BASE PAIRING IN 5-CHLOROURIDINE

By Charles L. Coulter and Stuart W. Hawkinson*

DEPARTMENT OF ANATOMY, UNIVERSITY OF CHICAGO

Communicated by George W. Beadle, June 9, 1969

Abstract.—5-Chlorouridine has been found to crystallize from water with the molecules of the nucleoside arranged in a base-paired, parallel-stranded ribbon. This type of polymer structure has not been previously considered for nucleic acids. We have constructed a model of polyuridylic acid based upon the 5-chlorouridine crystal structure and wish to suggest it as a plausible molecular complex for this and perhaps other polynucleotides.

The crystals are monoclinic, space group $P2_1$, with $a = 7.536$ A, $b = 5.790$ A, $c = 13.219$ A, and $\beta = 99.89°$. There are two 5-chlorouridine molecules per cell. The nucleoside bases are linked across a 2_1 axis with hydrogen bonds between N(3) and O(4). The hydrogen bond length is 2.85 A. The conformation of the base with the ribose about the β-glycosidic bond is *anti*, with a torsion angle relative to O (1′) of $-59.8°$. The sugar puckering is C(2′)-*endo*, and the conformation about the C(4′)—C(5′) bond is *gauche-gauche*. To build the polymer model, the C(4′)—C(5′) bond was rotated 100° and the sugar-base torsion angle adjusted to $-30°$. This brought O(3′) and O(5′) of adjacent sugars close enough to make the phosphate ester linkage found in polynucleotides.

Most of the homopolymers of the ribonucleotides have been found to interact to form ordered helical structures under appropriate conditions.[1-3] These ordered polymers have been studied extensively in solution[3, 4] and, using X-ray diffraction techniques, in the fibrous state. The results have been of considerable importance in understanding polynucleotide structure and behavior and in proposing models for RNA and DNA. Polyuridylic acid (poly U) has not been assigned a structure, although there is evidence for an ordered state at low temperatures.[5, 6] At temperatures above 10°C poly U does not show the marked hypochromism characteristic of polynucleotide complexes. From salt and temperature variation studies, Brown[6] has suggested that two structures are present under different conditions, one a multistranded complex and the other a single-stranded polymer. Simpkins and Richards[7] suggest that the pyrimidine bases in poly U are stacked in a parallel manner only in solutions of high salt concentration. Uracil can form satisfactory hydrogen-bonded base pairs[8] and does so in uracil crystals.[9]

As there is no clear reason why poly U cannot form an ordered complex, it seems possible in the light of our results that it forms a different type of two-stranded structure.

Materials and Methods.—5-Chlorouridine was purchased from the Calbiochem Corporation (Lot no. 680104, A grade). The compound had been crystallized from water. An unusually fine crystal was found in the material as supplied and was used for the X-ray diffraction analysis. The crystals are monoclinic, space group $P2_1$, with two molecules per cell. The crystallographic data are: $a = 7.536$ A, $b = 5.790$ A, $c = 13.219$ A, $\beta = 99.89°$, $d_{obs} = 1.63$ gm/cm^3, $d_{calc} = 1.626$ gm/cm^3.

Cell dimensions were derived by a least-squares fit of 31 careful 2θ measurements

and have a precision of ±0.002 A in the axes and ±0.02° in angle. X-ray diffraction data were collected on a General Electric XRD-5 diffractometer using CuK_α radiation. All data with $2\theta \leq 145°$ were collected; 1063 intensities were observed to have significant values, and 181 reflections were below the minimum observable intensity. The structure was solved by finding the two sets of chlorine to light atom vectors which occurred in the Patterson function. Three ribose atoms not definitively positioned from the Patterson vector map were located in a Fourier synthesis phased with the 15 atoms of known position. The structure was refined by least-squares methods using isotropic thermal parameters, and then using anisotropic thermal parameters. The 11 hydrogen atoms were placed on the basis of a difference Fourier synthesis, and refinement is continuing. The agreement factor, R, between the observed and calculated amplitudes is now 3.5%; the bond distances not involving hydrogen atoms have estimated standard deviations under 0.01 A, and the bond angles have an estimated error of ±0.5°. Positional parameters for the atoms are given in Table 1. Full details of the crystal structure analysis will be published at a later date. The structure appears to be isomorphous with the 5-bromouridine structure.[10]

Results.—The crystal structure of 5-chlorouridine is illustrated in Figure 1. The uracil bases are hydrogen-bonded across the 2_1 axis to neighbors $b/2$ above and below. The hydrogen bond length is 2.85 A, a value in excellent agreement with the 2.86 A found in uracil.[9] These are normal hydrogen bonds in both distance and orientation. The ribose conformation with respect to the base is *anti*, with a torsion angle,[11] ϕ_{CN}, of −59.8°. This is within the normal range found in nucleotides[12] and postulated for DNA.[11] The values found for cyclic uridine-3′,5′-phosphate[13] were −77° and −58°. The sugar puckering is C(2′)-*endo*, C(2′) being 0.54 A out of the least-squares plane through C(1′), C(3′), C(4′), and O(1′) on the same side as C(5′). Similar puckering was found for 5-fluoro-2′-deoxy-β-uridine[14] and for 5-bromouridine.[10] C(1′) of the ribose is 0.21 A out of the uracil plane, which corresponds to an 8.2° deviation of the glycosidic bond from the base plane. This is a feature common to many nucleotide structures; an angular deviation of 8° was also found for adenylic acid.[15] Bond distances and angles do not differ significantly from the expected values.

All atoms in 5-chlorouridine which can form hydrogen bonds do so, and the

TABLE 1. *Positional parameters of the atoms in 5-chlorouridine.*

Atom	x/a	y/b	z/c
Cl	0.1833	0.	0.0334
N(1)	0.1015	0.3576	−0.2262
C(2)	0.2236	0.5363	−0.2109
N(3)	0.3336	0.5346	−0.1156
C(4)	0.3304	0.3821	−0.0364
C(5)	0.2012	0.1991	−0.0612
C(6)	0.0920	0.1939	−0.1522
O(2)	0.2367	0.6846	−0.2740
O(4)	0.4308	0.4107	0.0464
C(1′)	−0.0392	0.3603	−0.3188
C(2′)	−0.0373	0.1475	−0.3862
C(3′)	−0.2366	0.1290	−0.4359
C(4′)	−0.3340	0.2145	−0.3514
C(5′)	−0.3942	0.0234	−0.2868
O(1′)	−0.2081	0.3618	−0.2861
O(2′)	0.0824	0.1856	−0.4560
O(3′)	−0.2772	0.2796	−0.5223
O(5′)	−0.2461	−0.1210	−0.2505

molecules are tightly packed in the crystal cell. In addition to the intermolecular links between the bases, O(3′) forms hydrogen bonds with the carbonyl oxygen, O(2), of a neighboring molecule and with the C(2′) hydroxyl group of a third molecule. The oxygen-oxygen distance in both cases is 2.81 A. The closest nonbonded contact distance is 3.05 A between the ribose O(1′) and an adjacent O(5′). This may be a weak hydrogen bond.

The immediate question of biochemical relevance is whether phosphate groups can be placed linking O(3′) and O(5′) of adjacent nucleosides. If so, the structure would be that of a polymer complex consisting of two parallel strands of poly U, base-paired to form an infinite ribbon. We have constructed such a model, and it is shown in Figure 2. The 2_1 screw axis present in the crystal was maintained in the preliminary model. This is not necessarily a requirement for the polymer, which might well have a slight twist within each strand. To build the polymer we had to bring the O(3′) and O(5′) atoms of adjacent sugars close enough to link them through a phosphate group. To this end, O(5′) was rotated

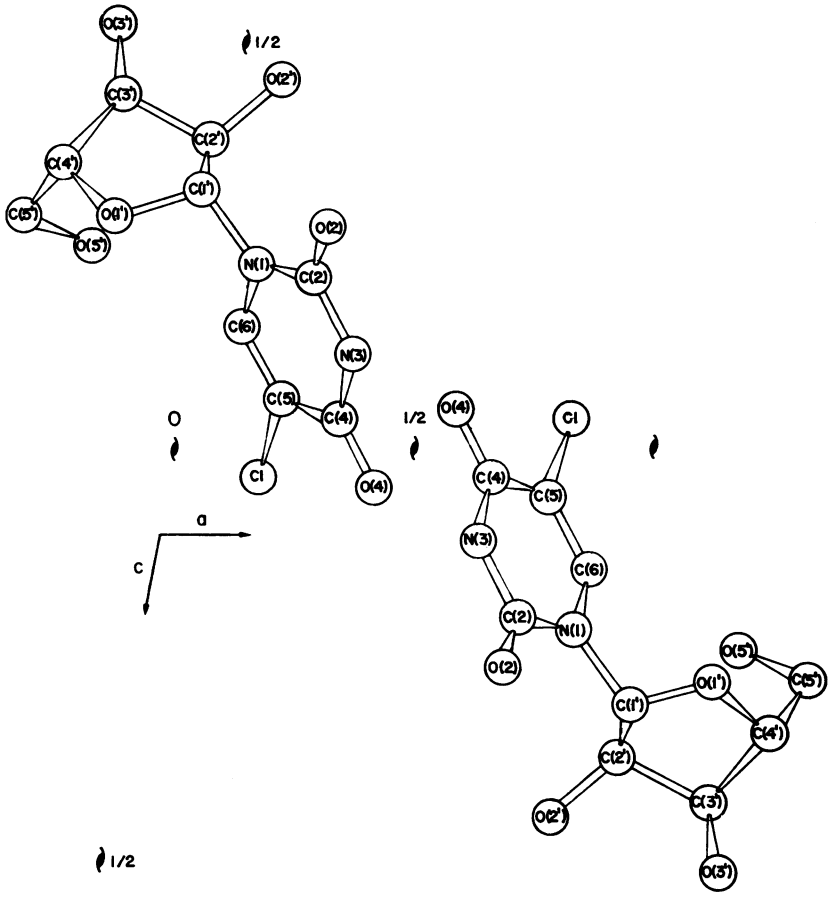

FIG. 1.—The crystal structure of 5-chlorouridine projected down the *b* axis.

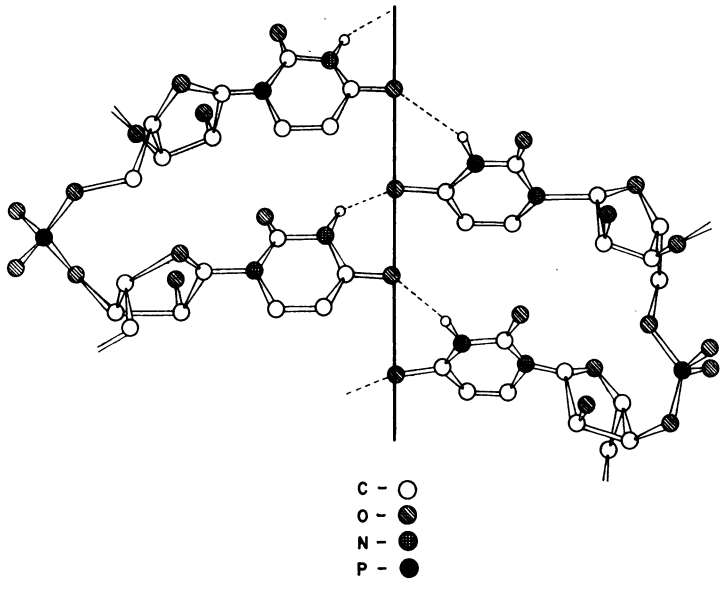

FIG. 2.—A schematic view of the proposed polymer structure. The polymer has a 2_1 axis relating the strands, with a vertical repeat of 5.8 A.

about the C(4′)—C(5′) bond by 100°, and the torsion angle of the base with respect to the sugar was adjusted to −30°. The conformation about the C(4′)—C(5′) bond was thus changed from *gauche-gauche* to *trans-gauche;* this is not the most frequently observed conformation,[16] but it has been found for uridine.[14] We also lengthened the C(5′)—O(5′) bond length from 1.41 A observed in 5-chlorouridine to 1.46 A, the value expected for a C(5′)—O(5′) bond in a phosphate ester.[13] A ϕ_{CN} of −30° is well within the normal *anti* range for sugar-base torsion angles.[12] These operations gave an O(3′)—O(5′) distance of 2.6 A, a value very near that found in phosphate esters.[13, 17] The position of the phosphate group was chosen so as to minimize nonbonded interactions. The conformation of the phosphorus atom about the C(3′)—O(3′) bond is *gauche-gauche* with respect to C(2′) and C(4′), and the dihedral angle between C(4′) and the phosphorus atom as viewed down the C(5′)—O(5′) bond is near 120°. This structure seems to meet the criteria for a plausible model.

Discussion.—The expected chemical and physical properties of this structure do not rule it out for poly U, particularly since the material was crystallized from water. Little or no hypochromism would be expected from such a polymer, since the bases are no longer stacked in a parallel fashion. Jakabhazy and Fleming[18] observed positive electric birefringence for poly U, indicating that the preferred base orientation is nearly parallel to the long axis of the polymer. The uracil bases are at an angle of 34° to the helix axis in our model. The viscosity of this material in solution would be expected to be lower than that found for DNA solutions or for solutions of DNA-like polymers, since the ribbons would be quite flexible. This structure is really intermediate between the doubly stranded

helices of the DNA type and a random coil. Transition between this structure and a parallel-stranded helical complex such as that proposed for acid poly A (ref. 1) would not be difficult. One hydrogen bond per base would be broken and reformed while the ribose-phosphate backbone twisted into a cylindrical helix. We would not necessarily expect sufficient orientation in gels of this material to permit well-oriented or crystalline fibers to be pulled. The structure is very open, and salt effects and the effects of other substances would be marked. The stabilization of the poly U structure by spermine[19] is an example of the latter phenomenon and may be compatible with these findings.

Other bases have been examined to see whether the occurrence of polymers of this type might be more generally expected. The C(1′)—C(1′) distance for a base pair in the crystal structure is 11.04 A, which is close to that found for Watson-Crick pairs.[3, 11] Thus, suitable arrangement of hydrogen bonding groups is the most stringent requirement. Cytosine can be substituted for uracil with no significant changes beyond rotation of the amino group on cytosine. Adenine fits less well, but there is a great deal of structural freedom here, and a similar model might well be built for poly A. Several of the 29 sterically feasible base pairs described by Donohue[8] and Donohue and Trueblood[11] would also fit in this scheme, and it is evident that this structure is less restrictive than the DNA structure in ruling out these possibilities. A fundamental question is whether this structure can be formed using antiparallel strands. If so, single-stranded nucleic acids might hydrogen-bond this way. The plausibility of such a structure is not immediately evident from our work. The obvious way to derive it from this model would be to rotate the sugars in one strand 180° about the glycosidic bond and then place the phosphate in a sterically acceptable position. This would leave the bases in a *syn* conformation with the ribose. Such conformations have been found,[12] but are unusual for nucleosides and nucleotides. Other possibilities are being explored.

We thank Mrs. Marsha Greaves for the figures and her assistance, and Dr. E. B. Fleischer for the use of his diffractometer. This research was supported by NSF grant GB-8103 by NIH grant FR-5367, and by the Dr. Wallace C. and Clara A. Abbott Memorial Fund of the University of Chicago.

* Public Health Service postdoctoral trainee under NIH-GM-94 training grant.
[1] Rich, A., D. R. Davies, F. H. C. Crick, and J. D. Watson, *J. Mol. Biol.*, **3**, 71 (1961).
[2] Langridge, R., and A. Rich, *Nature*, **198**, 725 (1963).
[3] Steiner, R. F., and R. F. Beers, *Polynucleotides* (Amsterdam: Elsevier, 1961), p. 235.
[4] Felsenfeld, G., and H. T. Miles, *Ann. Rev. Biochem.*, **36**, 407 (1967).
[5] Lipsett, M., these PROCEEDINGS, **46**, 445 (1960).
[6] Brown, R. A., *Arch. Biochem. Biophys.*, **115**, 102 (1966).
[7] Simpkins, H., and E. G. Richards, *Biopolymers*, **5**, 551 (1967).
[8] Donohue, J., these PROCEEDINGS, **42**, 60 (1956).
[9] Stewart, R. F., and L. H. Jensen, *Acta Cryst.*, **23**, 1102 (1967).
[10] Iball, J., C. H. Morgan, and H. R. Wilson, *Nature*, **209**, 1230 (1966); *Proc. Roy. Soc. London, Ser. A*, **295**, 320 (1967).
[11] Donohue, J., and K. N. Trueblood, *J. Mol. Biol.*, **2**, 363 (1960).
[12] Haschemeyer, A. E. V., and A. Rich, *J. Mol. Biol.*, **27**, 369 (1967).
[13] Coulter, C. L., *Acta Cryst.*, in press.
[14] Harris, R. D., and W. M. MacIntyre, *Biophys. J.*, **4**, 203 (1964).
[15] Kraut, J., and L. H. Jensen, *Acta Cryst.*, **16**, 79 (1963).

[16] Shefter, E., and K. N. Trueblood, *Acta Cryst.*, **18**, 1067 (1965).
[17] Sundaralingam, M., and L. H. Jensen, *J. Mol. Biol.*, **13**, 930 (1965).
[18] Jakabhazy, S. Z., and S. W. Fleming, *Biopolymers*, **4**, 793 (1966).
[19] Rogers, G. T., T. L. V. Ulbricht, and W. Szer, *Biochem. Biophys. Res. Commun.*, **27**, 372 (1967).